Genetics of Eosinophilic Esophagitis

Joseph D. Sherrill\textsuperscript{a} Carine Blanchard\textsuperscript{b}

\textsuperscript{a}Division of Allergy and Immunology, Department of Pediatrics, Cincinnati Children’s Hospital Medical Center, University of Cincinnati College of Medicine, Cincinnati, Ohio, USA; \textsuperscript{b}Nutrition and Health Research, Nestlé Research Center, Lausanne, Switzerland

Key Words
Eosinophilic esophagitis · Genetic disorder · Pathogenesis · Molecular pathways

Abstract
Eosinophilic esophagitis (EoE) is a complex genetic disorder characterized by eosinophilic inflammation within the esophagus. Multiple epidemiological studies estimate the prevalence of EoE is 4 in 10,000, with a higher disease prevalence in individuals of European ancestry and in males, highlighting a genetic etiology of the disease. EoE has often been noted to occur in multiple family members, particularly siblings, in a non-Mendelian pattern, indicating the heritable component of EoE is likely complex in nature. Although EoE is a newly diagnosed disorder involving a complex polygenic etiology, much progress has been made towards identifying the molecular pathways contributing to the disease pathogenesis and the genetic variants associated with disease susceptibility using a variety of approaches (genome-wide and candidate gene) as well as study designs (case-control and family-based cohorts). Here, we discuss the major scientific findings that have shaped the current molecular and genetic landscape of EoE as well as the major obstacles in the discovery of disease causal variants in complex disorders.

Genetic Predisposition

The high prevalence of eosinophilic esophagitis (EoE) among individuals of European descent and males noted across multiple epidemiological studies implicate a genetic contribution to the etiology of disease. Estimates for disease prevalence in these two populations have indicated nearly complete predominance in whites, while males comprise approximately three quarters of all EoE cases \cite{1-3}. For instance, in a study of 620 EoE patients collected over a period of 14 years, 90\% reported white ancestry and 75\% were males \cite{4}. Although reports differ as to whether disease manifestation and natural history are varied among races and between sexes \cite{5, 6}, much work remains to determine the genetic factors responsible for the increased disease prevalence in these at-risk populations.

In addition to gender and racial biases, EoE has often been noted to occur in multiple family members in a non-Mendelian pattern, indicating the heritable component of EoE is likely complex in nature. Indeed, a familial history of EoE in first-degree relatives has been reported in at least 6.8\% of EoE patients \cite{2}. Moreover, siblings of EoE patients are at substantially higher risk of developing disease, with a recurrence risk ratio of 80; these estimates suggest a 40-fold increase in disease risk when compared to siblings of asthmatic patients, where the sibling recurrence risk is 2 \cite{7}.
Molecular Pathogenesis of EoE

The histopathological hallmark of EoE is the presence of eosinophils in a hyperplastic esophageal epithelium. In 2001, Straumann et al. [8] revealed the Th2/allergic component of the disease by identifying IL-5-expressing T cells and IgE-bearing immune cells in esophageal biopsies of patients with EoE. The concept of allergic inflammation as a key initiating factor in EoE has been further supported in animal models demonstrating that eosinophil infiltration in the esophagus is associated with Th2 cytokines and mediators (e.g. IL-5, IL-13, and eotaxins) [8–17]. Indeed, human EoE is characterized by a variety of immune cells infiltrating the esophageal tissue, including eosinophils, mast cells, T cells, plasma cells, and dendritic cells [8, 9, 18, 19]. These works allowed the field to orient its efforts toward defining the precise allergic etiology of the disease, which has thus far proven particularly difficult. For instance, it has recently been shown that the development of EoE in an animal model is independent of IgE [20]. Additionally, in a subpopulation of EoE patients, no allergen or history of allergic diseases can be identified, suggesting again that multiple factors (both environmental and genetic) influence the disease susceptibility and development [2, 4, 21].

Global transcriptomic analysis of human esophageal biopsies from EoE patients and controls has shown more than 500 genes are dysregulated in the disease state compared to biopsies without any sign of disease [9]. The gene encoding the eosinophil chemoattractant eotaxin-3, chemokine (C-C motif) ligand 26 (CCL26), is the most upregulated gene in EoE [9]. Further in vitro work has demonstrated CCL26 expression can be induced by IL-13 treatment of esophageal epithelial cells. Indeed, IL-13 is highly increased in EoE and is responsible for many of the pathophysiological changes in the esophageal epithelium of EoE patients [9, 10]. In addition to CCL26, increased expression of mast cell and B cell genes and decreased expression of epidermal differentiation markers reflect the interplay between immune cells and the esophageal structural cells (i.e. epithelial cells and fibroblasts) in EoE pathogenesis [9]. Notably, these molecular pathways are relatively conserved between males and females, sporadic and familial EoE cases, pediatric and adult patients, and allergic and nonallergic EoE patients despite varying clinical manifestations [22, 23].

Periostin, a protein belonging to the fasciclin family, is highly upregulated in the lamina propria and the epithelium of the EoE patients. TGF-β and IL-13 are able to induce periostin expression in primary esophageal fibroblasts and, to a lesser extent, in primary esophageal epithelial cells. Using periostin-deficient mice, it was shown that periostin facilitates eosinophil recruitment into the esophagus [24]. Notably, in skin keratinocytes, periostin is able to induce thymic stromal lymphopoietin (TSLP), a potent cytokine in the initiation of allergic disease, thus suggesting a molecular loop between TGF-β, periostin, and TSLP [25]. TSLP has recently been shown to be key in the development and maintenance of EoE in an epicutaneous sensitization model. Indeed, TSLP is increased in the esophageal tissue of EoE patients [26, 27] and is associated with an increase number of basophils in the esophageal tissue [20]. These results highlight a complex network of cytokines (IL-13, TGF-β, periostin, and TSLP) and a previously unrecognized cell type (basophils) involved in EoE pathogenesis.

An increasing effort towards connecting the environmental triggers with the molecular pathogenesis of EoE has focused on epigenetic modifications. Epigenetics are heritable genomic marks that can be modulated in response to external stimuli to regulate gene expression independent of changes to the nucleotide sequence [28]. In EoE, histone acetylation at the eotaxin-3 promoter and altered expression of microRNAs (miRNA), which are short noncoding RNAs that regulate expression of target genes at the posttranscriptional level, have been observed in the esophageal tissue of EoE patients [29–31]. For instance, the miRNA miR-21 is upregulated in EoE and leads to decreased IL-12 production in T cells, potentially interfering with the Th1/Th2 balance [29]. Currently, there is an absence of prospective population-based epigenetic studies in at-risk individuals due to the relatively low prevalence of the disease. However, this type of cohort could ultimately allow the identification of the epigenetic modifications that predict the susceptibility to develop the disease.

Genetic Variants in EoE Susceptibility

A major step towards examining genetic susceptibility in EoE involved the unbiased screening of over 550,000 common single nucleotide variants (SNV) in a genome-wide association study of 351 EoE patients and 3,104 controls [27]. Surprisingly, a single locus on chromosome 5q22 met the genome-wide significance threshold (p < 5 × 10⁻⁸) and encompassed two genes: TSLP and WD repeat domain 36 (WDR36) [27]. Although variants in WDR36 had been previously associated with primary open-angle glaucoma [32] and, more notably, peripheral
blood eosinophilia [33], TSLP represented a more plausible candidate in EoE pathogenesis as alluded to previously (fig. 1). An epithelial-derived cytokine, TSLP, has been described as a master regulator of allergic inflammation that directly activates multiple immune cells [34–37]; in particular, TSLP-activated dendritic cells express elevated levels of OX40L and induce Th2 cytokine production in naïve T cells [38]. Esophageal expression of TSLP (but not WDR36) was elevated in patients with active EoE compared to healthy controls [27]. Moreover, the most significantly associated SNV in this study, rs3806932 (Fisher’s combined \( p = 3.19 \times 10^{-9} \), OR = 0.54–0.73), is located in the promoter region of TSLP and showed a genotype-dependent effect on TSLP expression in the esophageal tissue of patients with EoE [27].

Given the genetic [39] and biological [35–37] links implicating TSLP in asthma and atopic dermatitis, both prominent comorbidities in EoE [23, 40], the possibility for a spurious association with EoE due to confounding allergic phenotypes remained. However, using asthma as a covariate in a logistic regression analysis, TSLP variants retained a strong association with EoE risk [27]. The specificity of the genetic linkage between TSLP and EoE was further supported in a follow-up study in which genetic variants in epithelial-derived genes were genotyped in EoE patients, healthy controls, and atopic controls. Here, among the 738 SNVs genotyped in 53 candidate genes, TSLP was the most significantly associated genetic locus. Notably, when compared to atopic controls, the TSLP association was strengthened, whereas those initially seen with other Th2-related genes, such as IL4, were completely lost [26]. Furthermore, a nonsynonymous SNV (rs36133495) in the receptor for TSLP, cytokine receptor-like factor 2 (CRLF2), demonstrated a significant association within male EoE patients; this is of particular interest as CRLF2 is encoded on a pseudoautosomal region of the X and Y chromosomes.

### Table 1: Gene Expression and Disease Impact

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Gene symbol</th>
<th>Chromosome Associated SNP (Gene location)</th>
<th>Cellular source</th>
<th>Expression in EoE</th>
<th>Disease impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flaggrin</td>
<td>FG1</td>
<td>1q21.3 rs161816761 (exon)</td>
<td>Epithelial cells</td>
<td>↓</td>
<td>Reduced barrier, increased sensitization</td>
</tr>
<tr>
<td>WDR36</td>
<td></td>
<td>5p22.1 rs7723819 (promoter)</td>
<td>Unknown</td>
<td>Unchanged</td>
<td>Unknown</td>
</tr>
</tbody>
</table>

### Diagram 1: Molecular Pathogenesis of EoE

Fig. 1. The molecular pathogenesis of EoE.
chromosomes X and Y [41], which may play a contributory role in the male predominance of EoE.

The epithelial barrier gene filaggrin (FLG) represents yet another genetic locus linked with numerous allergic diseases, including EoE. The epithelium forms a protective barrier against environmental antigen exposure that, when compromised, can lead to antigen hypersensitivity and exacerbated immune responses [42]. Aggregation of FLG monomers is essential for the development of an intact stratum corneum, the superficial layer of the epithelium that is largely responsible for barrier formation [43, 44]. Loss-of-function variants in FLG have been identified in both asthmatic [45] and atopic dermatitis [46]. In a cohort of 365 EoE patients and 165 healthy controls, one FLG variant, 2282del4, demonstrated a significant association with disease (p = 0.017, OR = 5.0) even when restricted to a subset of nonatopic EoE patients (p = 0.03) [47].

Despite data implicating a heritable component in the genetics of EoE, the only family-based genetic study reported to date examined the prevalence of a noncoding SNV (rs2302009) in the 3′ untranslated CCL26 [9]. Transmission disequilibrium testing in trios, which consist of an affected offspring and heterozygous unaffected parents, demonstrated a biased inheritance of the G allele over the T allele in affected offspring (p = 5.4 × 10^-3, OR = 2.13) [9]. Notably, genotyping of rs2302009 in 117 EoE cases and 225 unrelated healthy controls also showed a significant association with disease (p = 6.9 × 10^-3, OR = 1.63) [9]. While the functional impact of this variant has yet to be determined, the normalization of CCL26 levels upon disease remission suggests a potential indirect effect on CCL26 expression. Several recent studies have shown SNVs in the 3′ UTR of genes can modify gene expression levels by altering their regulation by miRNAs [48–50]. These miRNA SNVs occur within miRNA-binding sites within target genes and can affect miRNA binding and thus the degree to which target miRNAs are regulated [51, 52].

In another candidate gene study examining corticosteroid responsiveness in EoE, a SNV in the promoter region of TGFB1 correlated with corticosteroid-induced disease remission. Here, homozygosity for a C to T transition at position −509 in the TGFB1 promoter was observed only in patients responding to corticosteroid treatment (n = 6 vs. n = 0 in nonresponsive patients) [53]. Mechanistically, the C-509T transition was shown to ablate a binding site for the transcription factor Yin Yang 1 in the TGFB1 promoter [54], thus suggesting lower TGF-β levels in corticosteroid responsive EoE patients. This hypothesis has been supported by other studies of the C-509T variant in asthma, where CC homozygous patients had significantly lower serum TGF-β levels compared to heterozygous and TT homozygous patients [55]. Although the cohort size (total n = 20) was comparatively small with respect to most genetic studies to infer disease causality, the substantial biological evidence implicating TGF-β in EoE pathogenesis warrants further investigation into TGFB1 risk variants in EoE. In particular, esophageal mastocytosis and elevated expression of TGFB1 and mast cell-specific genes including carboxypeptidase A3 (CPA3) within the inflamed esophageal mucosa have been noted in EoE [56–58]. Notably, mast cells represent a significant cellular source for TGF-β, which can directly stimulate esophageal smooth muscle contraction in vitro [59]. Moreover, TGF-β is implicated in mediating the chronic pathobiological symptoms of EoE as TGF-β can directly promote epithelial to mesenchymal transition [60].

The interest in disease risk variants affecting the TGF-β pathway has been further heightened by the recent observation of the coexistence of EoE with inherited connective tissue disorders (CTD). An eightfold risk of EoE was identified in patients with a CTD [61]. The CTD spectrum includes those with clinically defined genetic etiologies such as Marfan syndrome, Ehlers-Danlos syndrome, and Loeys-Dietz syndrome. These disorders have disease-specific variants in TGF-β interacting proteins like fibrillin 1 (FBN1; Marfan syndrome) [62], collagen V (COL5A1 and COL5A2; Ehlers-Danlos syndrome) [63], TGFB2 (Loeys-Dietz syndrome) [64], and TGF-β receptor type I or II (TGFBRI and TGFBRII; Loeys-Dietz syndrome and Ehlers-Danlos syndrome) [65, 66]. These findings strongly suggest genetic dysregulation of the TGF-β pathway underlies disease modalities that are germane to both EoE and CTDs.

**Challenges to EoE Genetic Studies**

A major hurdle to the advancement of the genetic underpinnings of EoE stem from the modern recognition of EoE as a bona fide clinicopathologic disorder. Consensus clinical recommendations for the diagnosis and treatment of EoE were originally published in 2007 [67] and then updated in 2011 [68]. Coupled with the rarity of the condition, this has precluded the ascertainment of large case cohorts, effectively reducing statistical power and hampering the much-needed meta-analyses for common risk variants in complex diseases. Combining multiple cohorts has been quite effective in genome-wide association studies for inflammatory bowel disease risk variants,
where analyses of over 15,000 cases have nearly doubled the number of disease risk loci identified [69, 70]. Moreover, the difficulty (and differences in clinical practices) in distinguishing EoE from other related disorders, namely gastroesophageal reflux disease, could hamper large population genetic studies wherein the success relies on a homogenous, well-phenotyped case cohort. In addition, as several subphenotypes [71] and comorbidities [61, 72] in EoE have emerged, identifying the true genetic causality in EoE has been further complicated.

The recent clinical acceptance of EoE not only limits the number of cases available for genetic studies, but also the age of affected individuals. This poses a particular hindrance to familial studies of multiple affected generations with clinically diagnosed EoE. Instead, such studies are often relegated to rely upon patient self-reported EoE-like symptoms, particularly in older relatives who are either unlikely to undergo an esophagogastroduodenoscopy or are deceased. Interestingly, retrospective studies have shown esophageal eosinophilia with concomitant atopic phenotypes indicative of EoE in patients originally diagnosed with reflux esophagitis dating back to 1982, thus supporting a historically misdiagnosed (or undiagnosed) cohort of potential EoE cases [73, 74].

Additional challenges to interpreting genetic studies in EoE and other complex diseases relates to the broad spectrum of presenting phenotypes. For such polygenic diseases, penetrance, which measures the presence of a disease phenotype in a population of known genotypes at a given locus, is often incomplete [75]. This is exemplified in the TSLP/WDR36 association with EoE, where the risk variants are present in 30–40% of the healthy unaffected control populations and the causal variant(s) remains to be identified [26, 27]. Moreover, expressivity, or the degree to which a given disease phenotype is expressed in individuals with identical genotypes, further complicates identification of causal genetic variants [75]. For instance, the patchiness of esophageal inflammation and diagnostic threshold of 15 eosinophils per high-power field in EoE could preclude patients with persistent low-grade esophageal inflammation (e.g. ≤14 eosinophils per high-power field) with endoscopic findings and disease symptoms characteristic of EoE [76].

### Next Challenges

Despite the rapid progress towards elucidating the molecular pathogenesis of EoE, much work remains. In particular, the identification of causal variants has been hindered due, in part, to the technical and statistical limitations of population-based studies, which require large cohort sizes and have thus far focused on common variants with small effects. However, the expansion of genome-wide SNV genotyping chips and technological advances in large-scale (whole-genome or whole-exome) DNA sequencing allow for the interrogation of rare SNVs (minor allele frequencies <1%). Moreover, as EoE becomes increasingly diagnosed, the rise in the numbers of highly penetrant, multiplex families will yield an increasing number of family-based genetic studies with the promise of identifying causal variants that are inherited among the affected members. While these variants may be family-specific, it is conceivable that commonly affected biological pathways will begin to emerge across multiple affected families and unrelated patients.

### Disclosure Statement

The authors declare that no financial or other conflict of interest exists in relation to the content of the article. C.B. is employed by NESTEC Ltd.

---

**References**

Genetics of EoE


